מטלת מנחה 15 - קורס 20425

שאלה 1

א. נמצא אח"ה ל p. על פי לינאריות התוחלת:

$$E[\sum_{i=1}^{m} X_i] = \sum_{i=1}^{m} E[X_i] = \sum_{i=1}^{m} n \cdot ip = np \cdot \sum_{i=1}^{m} i = np \cdot \frac{m(m+1)}{2}$$

על כן אם ניקח את האומד $\hat{p} = \frac{2}{nm(m+1)}\Sigma X_i$ נקבל:

$$E[\hat{p}] = \frac{2}{nm(m+1)} \cdot E[\sum_{i=1}^{m} X_i] = \frac{2}{nm(m+1)} \cdot np \cdot \frac{m(m+1)}{2} = p$$

ב. נקבל ע"פ 9.15:

$$MSE(\hat{p}, p) = Var(\hat{p}) + (E[\hat{p}] - p)^2 = Var(\hat{p}) + 0^2 = Var(\hat{p})$$

כאשר:

$$Var(\hat{p}) = Var(\frac{2}{nm(m+1)}\Sigma X_i) = (\frac{2}{nm(m+1)})^2 \cdot Var(\Sigma X_i)$$

המשתנים המקריים X_i בלתי תלויים, לכן:

$$Var(\sum_{i=1}^{m} X_i) = \sum_{i=1}^{m} Var(X_i) = \sum_{i=1}^{m} nip(1 - ip) = np\sum_{i=1}^{m} i - pi^2 =$$

$$= np\sum_{i=1}^{m} i - np^2\sum_{i=1}^{m} i^2 = np\frac{m(m+1)}{2} - np^2\frac{m(m+1)(m+2)}{6} =$$

$$= \frac{1}{6}npm(m + 1)(3 - p(m + 2))$$

ומקבלים:

$$MSE(\hat{p}, p) = (\frac{2}{nm(m+1)})^2 \cdot \frac{1}{6}npm(m + 1)(3 - p(m + 2)) =$$

$$= \frac{4}{n^2m^2(m+1)^2} \cdot \frac{nm(m+1)}{6} \cdot p(3 - p(m + 2)) =$$

$$= \frac{2p(3-p(m+2))}{3nm(m+1)}$$

ג. נקבל ע"פ 9.13:

$$(1 \hat{-} p) = 1 - \hat{p} = 1 - \frac{2}{nm(m+1)}\Sigma X_i$$

ד. כמו כן מקבלים:

$$MSE(1 \hat{-} p) = Var(1 \hat{-} p) + 0^2 = Var(1 - \hat{p}) = (-1)^2Var(\hat{p}) =$$

$$= \frac{2p(3-p(m+2))}{3nm(m+1)}$$

שאלה 2

יהא $X_1, X_2,..., X_n$ מדגם מקרי של ההתפלגות $X \sim NB(r, p)$. פונקציית ההתפלגות המשותפת של המדגם תהיה, לפי אי-תלות חלקי המדגם,

$$p(x_1, x_2,..., x_n) = \prod_{i=1}^{n} P\{X_i = x_i\} = \prod_{i=1}^{n} \binom{i-1}{r-1} (1 - p)^{i-r} p^r$$

לכן פונקציית הנראות תהיה:

$$L(p) = \prod_{i=1}^{n} \binom{x_i-1}{r-1} (1 - p)^{x_i-r} p^r$$

וכן נקבל:

$$\ln[L(p)] = \ln(\prod_{i=1}^{n} \binom{x_i-1}{r-1} (1 - p)^{x_i-r} p^r) = \sum_{i=1}^{n} \ln\binom{x_i-1}{r-1} (1 - p)^{x_i-r} p^r =$$

$$= \sum_{i=1}^{n} \ln\binom{x_i-1}{r-1} + \sum_{i=1}^{n} \ln(1 - p)^{x_i-r} + \sum_{i=1}^{n} \ln p^r =$$

$$= \sum_{i=1}^{n} \ln\binom{x_i-1}{r-1} + \ln(1 - p) \cdot \sum_{i=1}^{n} (x_i - r) + \ln p \cdot \sum_{i=1}^{n} r =$$

$$= \sum_{i=1}^{n} \ln\binom{x_i-1}{r-1} + (\sum_{i=1}^{n} x_i - nr) + \ln(1 - p) + nr \cdot \ln p =$$

נרצה למצוא אומדן נראות מקסימלית. ידוע על פי מונוטוניות פונקציית ה ln כי $L'(p) = 0,\ L''(p) < 0$ אם ורק אם $\ln[L(p)]' = 0,\ \ln[L(p)]'' < 0$. לכן נגזור את $\ln[L(p)]$ לפי p:

$$\ln[L(p)]' = (\sum_{i=1}^{n} x_i - nr) \cdot \frac{-1}{1-p} + nr \cdot \frac{1}{p} = 0$$

נכפול ב $p(1 - p)$ ונקבל:

$$-(\Sigma x_i - nr)p + nr(1 - p) = 0$$

$$nr \cdot p - \Sigma x_i \cdot p + nr - nr \cdot p = 0$$

$$nr = \Sigma x_i \cdot p$$

$$\hat{p} = \frac{nr}{\Sigma x_i}$$

נוודא מקסימליות:

$$\ln[L(p)]'' = (\sum_{i=1}^{n} x_i - nr) \cdot \frac{-1}{(1-p)^2} + nr \cdot \frac{1}{p^2} = -\frac{\sum_{i=1}^{n} x_i - nr}{(1-p)^2} - \frac{nr}{p^2} < 0$$

הערה: Σx_i מציין את מספר ה"נסיונות" הכולל עד במדגם המקרי, והוא לפחות n פעמים r נסיונות.

ב. עבור $n = 10, r = 3$:

$$P\{X = 8\} = \binom{7}{2} (1 - p)^5 p^3$$

על כן נקבל לפי 9.20 את האנ"מ:

$$\hat{P} = 21(1 - \frac{30}{\Sigma x_i})^5 (\frac{30}{\Sigma x_i})^3$$

כמו כן, מתקיים $E[X] = \frac{r}{p}$, ולכן אומד יהיה

$$\hat{\mu} = \frac{r}{\frac{nr}{\Sigma x_i}} = \frac{\Sigma x_i}{n} = \bar{X}$$

שאלה 3

א. ננסח את הבעיה כבעיה של בדיקת השערות. נסמן ב p_0 את התפלגות החבילות ביקב "הילולה", וב p_1 את התפלגות החבילות ביקב "השומרים".
נסמן ב X_i את מספר בקבוקי היין הלבן בחבילה ה i. המשתנה שאנחנו בודקים: $X = X_1 + X_2$ מספר בקבוקי היין הלבן הכולל. ההתפלגויות:

מספר בקבוקי יין לבן	0	1	2
"הילולה"	0.1	0.2	0.7
"השומרים"	0.4	0.3	0.3

השערת האפס: $X_i \sim p_0$, כלומר, החבילות הגיעו מיקב "הילולה"
השערת המחקר: $X_i \sim p_1$, כלומר החבילות הגיעו מיקב "השומרים"

נחשב את ההסתברויות לקבלת $X = x$ ב-2 ההתפלגויות בעזרת נוסחת ההסתברות השלמה, וכן נחשב את יחס הנראות:

$X = x$	אפשרויות ל (X_1, X_2)	$p_0(X = x)$	$p_1(X = x)$	$\lambda(x)$
0	(0,0)	$0.1^2 = 0.01$	$0.4^2 = 0.16$	16
1	(0,1), (1,0)	$2 \cdot 0.2 \cdot 0.1 = 0.04$	$2 \cdot 0.4 \cdot 0.3 = 0.24$	6
2	(2,0), (1,1), (0,2)	$2 \cdot 0.7 \cdot 0.1 + 0.2^2 = = 0.18$	$2 \cdot 0.4 \cdot 0.3 + 0.3^2 = 0.33$	1.833
3	(2,1), (1,2)	$2 \cdot 0.7 \cdot 0.2 = 0.28$	$2 \cdot 0.3 \cdot 0.3 = 0.18$	0.6429
4	(2,2)	$0.7^2 = 0.49$	$0.3^2 = 0.09$	0.1837

נבנה מרחב דחייה C אליו נוסיף בסדר יורד את הערכים אשר $\lambda(x)$ מרבי עבורם, כל עוד נקבל שרמת המובהקות לא עולה על 5%:

C	$\alpha = p_0(C)$
$\{0\}$	0.01
$\{0, 1\}$	0.05
$\{0, 1, 2\}$	0.23

נקבל שמרחב הדחייה יהיה $C = \{X = 0, 1\}$ כאשר X מייצג, כאמור, את מספר בקבוקי היין הלבן בשתי החבילות יחד. המבחן הוא בעל רמת מובהקות של 5%, ועוצמה מקסימלית עבור רמת מובהקות זו על פי ניימן-פירסון.

ב. ההסתברות לטעות מסוג 1 תהיה $\alpha = p_0(C) = 0.05$

ההסתברות לטעות מסוג 2 תהיה $\beta = p_1(\bar{C}) = p_1(X = 2, 3, 4) = 0.33 + 0.18 + 0.09 = 0.6$

ג. אם נוסיף חסם $\alpha \leq 0.1$ המבחן לא ישתנה, שכן לא נוכל להוסיף עוד איברים למרחב הדחייה מבלי לעלות על רמת מובהקות של 10%. האיבר הבא בתור להוספה למרחב הדחייה ייתן רמת מובהקות של 23%.